

FIRST-ORDER FRAMEWORKS FOR MANAGING MODELS IN ENGINEERING OPTIMIZATION

NATALIA M. ALEXANDROV*, ROBERT MICHAEL LEWIS†

Abstract

Approximation/model management optimization (AMMO) is a rigorous methodology for attaining solutions of high-fidelity optimization problems with minimal expense in high-fidelity function and derivative evaluation. First-order AMMO frameworks allow for a wide variety of models and underlying optimization algorithms. Recent demonstrations with aerodynamic optimization achieved three-fold savings in terms of high-fidelity function and derivative evaluation in the case of variable-resolution models and five-fold savings in the case of variable-fidelity physics models. The savings are problem dependent but certain trends are beginning to emerge. We give an overview of the first-order frameworks, current computational results, and an idea of the scope of the first-order framework applicability.

Key Words: Approximation concepts, approximation management, model management, nonlinear programming, optimal design, surrogate optimization

1 Introduction

Computational models in science and engineering have progressed steadily in numerical accuracy and physical fidelity as computers have become more powerful. It is, arguably, to be expected that computational models will grow in complexity to consume all computing power available at any given moment. This makes the objectives of an analyst at computational odds with the objectives of a designer. Given a set of design or control variables x , the analyst wishes to solve, as accurately as possible, a disciplinary analysis equation (or a system of coupled disciplinary analysis equations, in the case of multidisciplinary optimization [4, 38])

$$A(x, u(x)) = 0 \tag{1}$$

for state variables u that describe the physical behavior of the system. System (1) is frequently an expensive simulation, say, involving the solution of a set of differential equations, as in the case of computing the flow around an airplane. The designer, on the other hand, wishes to solve a problem of the form

$$\begin{aligned} & \underset{x}{\text{minimize}} && f(x, u(x)) \\ & \text{subject to} && h(x, u(x)) = 0 \\ & && g(x, u(x)) \geq 0 \\ & && x_l \leq x \leq x_u, \end{aligned} \tag{2}$$

where, given x , $u(x)$ is determined from (1). Unfortunately, the use of high-fidelity models, such as those based on fine computational meshes (high-resolution) or detailed physics (e.g., the Navier-Stokes equations) at every iteration of an optimization procedure can be prohibitively expensive.

*Multidisciplinary Optimization Branch, MS 159, NASA Langley Research Center, Hampton, VA, 23681-2199, n.alexandrov@larc.nasa.gov.

†Department of Mathematics, College of William & Mary, P.O. Box 8795, Williamsburg, VA, 23187-8795, buckaroo@math.wm.edu. This author's research was supported by the National Aeronautics and Space Administration under NASA Contract No. NAS1-97046 while the author was in residence at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA 23681.

The difference between the objectives of the analyst and the designer suggests that the designer can be satisfied with intermediate results of lower physical fidelity, provided they lead to improved designs. Only at the optimal solution must the design be consistent with the accurate, high-fidelity analysis. Given this distinction, researchers have long taken advantage of computational approximations and models of varying accuracy in engineering design optimization[†]. A survey on the use of approximations in structural optimization can be found in [11], while recent overviews of methods for aerodynamic analysis and optimization can be found in [25, 34].

Until recently, procedures for using variable-fidelity approximations and models in systematic optimization had been largely based on heuristics, and convergence to a solution of the highest-fidelity optimal design problem had not been guaranteed, in general. With a few exceptions [10, 32, 35], the analysis of the use of approximations in optimization had focused on the question of convergence to a solution of the approximate problem (e.g., [16, 24]). Due to improvements in numerical modeling techniques and the increased availability of high-fidelity analyses, optimization with variable-fidelity approximations has become a subject of much interest in the past few years (e.g., [17, 23]).

This paper discusses an approach, approximation/model management optimization (AMMO), that facilitates design optimization, using approximations and models in systematic ways. This approach (previously presented in [2, 3, 6, 7, 29]) alleviates the expense of relying exclusively on high-fidelity models by incorporating well-established engineering approximation concepts with ideas from nonlinear programming that ensure global convergence of the overall process and robust performance. We focus here on an overview of our experience with the practical effectiveness of the methods originally proposed in [2, 3, 29].

The paper is organized as follows. The next section looks at the ideas that underlie the first-order AMMO methods. The following section briefly describes the types of lower-fidelity models used and the consistency conditions we impose via corrections. The computational demonstrations are then described. The paper concludes with lessons learned to date and some mention of ongoing work.

2 First-order AMMO frameworks

In conventional optimization, the analysis (1) supplies the optimizer with objective and constraint function and derivative information while the optimizer produces new values of the design variables x for re-analysis. The optimizer uses the function and derivative information to build local approximations—usually first or second-order Taylor series. If evaluating the problem functions and derivatives involves a simulation of high accuracy but high computational cost (e.g., the Navier-Stokes equations), the repeated analyses required by the optimizer are expensive.

The basic idea of first-order model management is to replace the local Taylor series model in the optimization subproblems with more general approximations or surrogates. These surrogates are not arbitrary but satisfy certain first-order consistency conditions with respect to the high-fidelity model. The consistency conditions ensure that the overall optimization process based on the less expensive models converges to a solution of the high-fidelity optimization problem. We occasionally have recourse to the expensive, high-fidelity computations to verify that we are indeed generating improved designs, and also to re-calibrate the lower-fidelity model, based on a set of systematic criteria.

[†]Some authors make distinctions in the use of the terms “models”, “surrogates”, and “approximations”. For simplicity, the terms are used interchangeably here.

AMMO is based on the trust-region idea [31,33] in nonlinear programming, which can be understood as an adaptive move limit strategy for improving the global behavior of optimization algorithms based on local models. The trust-region methodology ensures the convergence of the AMMO scheme to a solution of the higher-fidelity problem by providing a measure of the surrogate's predictive behavior, a criterion for updating the surrogate, and a systematic response to situations in which an optimization phase performed using a surrogate gives either an incorrect or a poor prediction of the higher-fidelity model's actual behavior.

3 Consistency, convergence, and performance

The first-order model management idea can be used in conjunction with any gradient-based optimization algorithm. See [5,6] for a discussion of several such algorithms. Here we give an example of an AMMO algorithm for bound constrained minimization.

```

Initialize  $x_c, \Delta_c$ 
Do until convergence:
  Select model  $a_c$  with  $a_c(x_c) = f(x_c); \nabla a_c(x_c) = \nabla f(x_c)$ 
  Solve approximately for  $s_c = x - x_c$ :
    minimize $s$   $a_c(x_c + s)$ 
    subject to  $x_l \leq x \leq x_u$ 
                $\|s\|_\infty \leq \Delta_c$ 

  Compute  $\rho_c \equiv \frac{f(x_c) - f(x_c + s_c)}{f(x_c) - a_c(x_c + s_c)}$ 
  Accept  $s_c$  if  $f(x_c) > f(x_c + s_c)$ ; otherwise reject
  Update  $\Delta_c$ 
End do

```

The trust radius Δ_c is updated according to the usual trust region practice [31]. The hope is that the bulk of the computational expense will involve calculations based on the less expensive approximation a_c .

The first-order consistency conditions in this case require the approximation a_c used in the optimization subproblem to satisfy

$$a_c(x_c) = f(x_c) \tag{3}$$

$$\nabla a_c(x_c) = \nabla f(x_c). \tag{4}$$

These ensure that a_c mimics the local behavior of a Taylor series model around our current best design x_c . This, in turn, can be used to prove that the overall optimization process will converge to a constrained stationary point of the high-fidelity objective f [5].

A number of corrections can be used to easily ensure the conditions (3)–(4). In the work reported here, we use a correction technique called the β -correlation method, due to Chang et al. [19]. Given the high-fidelity objective $f_{hi} = f$ and any low-fidelity approximation f_{lo} of the objective f_{hi} , we correct f_{lo} as follows. Define

$$\beta(x) = \frac{f_{hi}(x)}{f_{lo}(x)}$$

and construct

$$\beta_c(x) = \beta(x_c) + \nabla \beta(x_c)^T (x - x_c).$$

Then

$$a_c(x) = \beta_c(x)f_{I_0}(x)$$

satisfies the consistency conditions (3)–(4). For an alternative, additive correction scheme, see [30].

Convergence analysis of the resulting AMMO schemes relies on the consistency conditions and standard assumptions for the convergence analysis of the underlying optimization algorithm [5]. For general problems, our current preferred AMMO scheme is based on sequential quadratic programming.

Practical efficiency of any particular AMMO scheme is a separate issue. The ability to transfer the computational load onto the lower-fidelity, cheaper computations, and thereby reduce the overall computational cost, will depend on the predictive qualities of the surrogates. Note that even though the surrogate models may not be good approximators of the higher-fidelity models for the purposes of analysis, they may possess suitable predictive properties for the purposes of optimization. That is, an approximation may not capture all the important properties of a higher-fidelity function, but it may still produce a step that will lead to a satisfactory improvement in the merit function for the higher-fidelity problem. The computational demonstrations we present in Section 5 validate the effectiveness of AMMO.

4 Origins of low-fidelity models

There exist many possible ways to construct low-fidelity models that we can then correct and use inside the AMMO framework. Engineers have frequently turned to models based on fitting surfaces to samples of high-fidelity responses, such as classical polynomial response surface models, splines, or, more recently, kriging models. Data-fitting approximations are attractive because they rely directly on high-fidelity data and do not, in general, require sensitivity information. These models are also generally simple to construct. However, such models suffer from a number of well-known limitations. These include the question of constructing an appropriate sampling scheme and obtaining a sufficient sampling of the design space. These problems become worse as the number of design variables increases—the so-called “curse of dimensionality”. A more subtle difficulty in constructing such models is discussed in Section 6.

Variable-fidelity approximations that avoid the curse of dimensionality present an attractive alternative to approximations based on data-fitting. We are considering a number of options. Variable-accuracy models arise when a particular analysis can be converged to a user-specified tolerance. For instance, aerodynamic analysis codes are usually equipped with a user-specified convergence criterion. Another mechanism for managing variable-accuracy models can be found in [18].

Variable-resolution models are computed by executing a single physical model on meshes of varying degree of refinement. Such models are frequently available in engineering optimization, and we discuss one instance later. An elaboration of the fundamental AMMO idea leads to the multigrid optimization scheme discussed in [30].

Finally, the most provocative choice of models is that of variable-fidelity physics models. For instance, in aerodynamics, the physical models range from linear potential models that describe inviscid, irrotational, incompressible flow to Navier-Stokes equations for nonlinear viscous flow. Our demonstrations include such a test case.

5 Computational demonstrations

Because of the limitations of approximations based on data fitting discussed in the previous section, we have focused on models that are independent of the number of variables: variable-resolution models and variable-fidelity physics models. Independence of dimensionality is of special importance, because in preliminary design, multidisciplinary applications of even a modest size may number a few hundred variables [40, 41]. Aerodynamic shape optimization—one of the most computationally intensive single-discipline problems of interest in aerospace design—is also of a dimensionality that defies current data-fitting approximation techniques.

As we discussed in the previous sections, first-order AMMO frameworks admit a wide variety of models. Because current convergence results are of a global nature, extensive computational experience has to be accumulated to validate the actual performance of AMMO, in terms of the savings in high-fidelity function and derivative evaluations. The demonstrations of this section are aimed at accumulating such realistic computational experience.

Computational experiments with all AMMO frameworks and all models are done as follows. The problems are solved with high-fidelity function and derivative evaluations, using well-known commercial optimization software[§], such as NPSOL [21] and PORT [26], in order to obtain a baseline number of function and derivative evaluations or iterations to find an optimum. The problems are then solved in the AMMO framework under investigation.

Because it is difficult or impossible to predict *a priori* the relative descent characteristics of a model, the computational tests include cases when the relationship between the various levels of models is favorable and the cases when it is not. Qualitatively, the relationship is favorable if the lower-fidelity model provides a long sequence of descent steps for the higher-fidelity merit function before the lower-fidelity model has to be re-calibrated. The relationship is not favorable when the lower-fidelity model does not satisfactorily capture the trends in the high-fidelity objective and constraints on a significant region of the feasible region. In fact, as we discuss later, the variable-fidelity physics demonstration has proved to be a test of the worst-case scenario. In the remainder of this section, we present the numerical results of demonstration problems for two types of models.

5.1 Managing variable-resolution models

The use of a single physical model evaluated on meshes of a varying degree of refinement appeared as a first natural step toward managing variable-fidelity models not based on data-fitting techniques. As we discuss, in our computational demonstration of AMMO on two problems of aerodynamic optimization, the suite of variable-resolution models was provided by the Euler equations solved on a variety of meshes. The finer the mesh, the higher the model fidelity and computational expense. We anticipated reasonable savings with model management because the responses of interest associated with the variable-resolution models may be expected to have similar global trends.

Initially, the meshes constructed for the models were arbitrary meshes of different size, with no relation among them. While sufficiently fine meshes should, in principle, produce consistent functions, the meshes were too coarse to observe this effect. Instead, objectives and constraints computed on unrelated meshes had wildly disparate trends and features, a phenomenon noted by other investigators [14]. And, given the relative coarseness of the meshes, it was not even clear that the finer of the meshes produced the higher-fidelity result. The difficulty was remedied by using coarser meshes that were proper subsets of the finest mesh.

[§]The use of names of commercial software in this paper is for accurate reporting and does not constitute an official endorsement, either expressed or implied, of such products by the National Aeronautics and Space Administration.

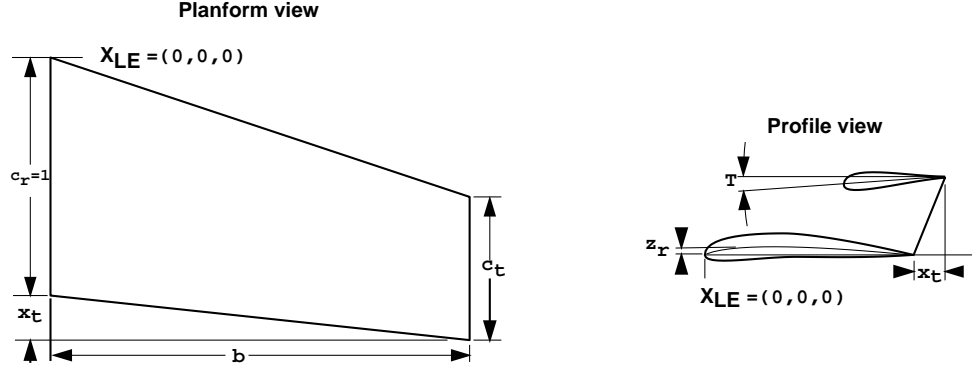


Figure 1: The 3D wing problem

Three-dimensional aerodynamic wing optimization is the first demonstration problem. The wing consists of a single trapezoidal panel with a rounded tip. It is parameterized by fifteen variables. The wing and some of the associated parameters are depicted in Figure 1. We chose to use two independent variables—the tip chord and the tip trailing-edge setback—in the initial demonstration to ease the visualization of the results. The objective function is the negative lift-to-drag coefficient ratio

$$f(x) = -C_L/C_D.$$

Several artificial constraints simulate multidisciplinary constraints:

1. A lower bound on total lift $C_L \times S$ simulates a minimum payload requirement, where S is the semispan wing planform area;
2. An upper bound on C_M (pitching moment coefficient) simulates a trim constraint;
3. An upper bound on C_l (rolling moment coefficient) simulates a maximum bending moment constraint.

The aerodynamic analysis code used for this study is CFL3D.ADII [39], a version of CFL3D [36] obtained via the ADIFOR automatic differentiation tool [12]. The surface geometry was computed based on the problem parameters via software that uses the RAPID technique [37]. The ADIFOR generated analysis code includes the capability for computing the gradients. The volume mesh and associated gradients needed for CFL3D are generated using an ADIC [13] generated version of CSCMDO. Here we consider the case of two variable-fidelity models and associated constraints, generated by performing the CFL3D.ADII analysis on a low-fidelity mesh $97 \times 25 \times 17$ and a high-fidelity mesh $193 \times 49 \times 33$.

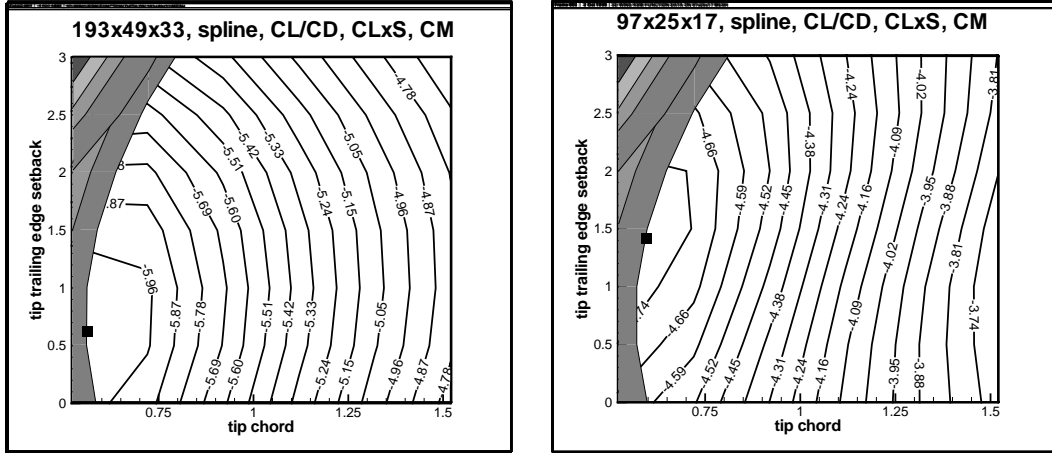
Since the analysis is based on a multigrid solution process, the CPU time per converged function evaluation is linear in the number of grid points, resulting in an eight-fold difference in execution time between adjacent levels of fidelity. For instance, on an Ultra 1 Sun workstation, a single function and constraint evaluation on the $97 \times 25 \times 17$ mesh takes eight minutes, and the $193 \times 49 \times 33$ mesh analysis takes about an hour, without computing derivatives.

Initial tests were conducted with the actual function evaluations obtained by executing the analysis software. An examination of the problem functions revealed that they exhibited benign behavior in that the objectives and constraints were smooth and very nearly convex. Figure 2 depicts the level sets of the objective functions and active constraints obtained by performing analyses on the $193 \times 49 \times 33$ and $97 \times 25 \times 17$ meshes. The shaded regions are infeasible. For

the subsonic case under study, constraint C_l is inactive and is not depicted. Black squares mark the solutions. Note that this problem has a favorable structure for AMMO. Although the optima are at different locations, the low-fidelity and high-fidelity objectives and constraints have similar trends.

Thus, the difficulty of experimenting with the actual analyses was due to the computational expense of obtaining function and gradient information. Because the actual functions were benign computationally, we made a decision to expedite the numerical experimentation by using substitutes based on data fitting to conduct the tests for this problem. The substitutes were not used in the conventional way, i.e., to approximate the lower-fidelity model. Instead, we used the substitutes for both the high-fidelity model and the low-fidelity model to obtain an estimate of AMMO performance at a small fraction of the computational expense. Two-dimensional, uniform, variation diminishing splines (obtained with the help of the PORT [26] package) provided the best approximation to the actual functions and were taken to represent the “true” high-fidelity and low-fidelity models.

An additional benefit of using substitutes in this series of tests became apparent quickly. As Figure 2 indicates, for this problem, the lower-fidelity functions obtained on coarser meshes provide an excellent approximation (in regards to the general trends) to those computed on finer meshes. Because this is the most favorable scenario for AMMO, one must also investigate cases where the lower-fidelity problem does not capture the high-fidelity descent behavior well. Substitutes provided by kriging, implemented locally, gave only a slightly less accurate approximation of the actual functions than the splines (see Figure 3), thus retaining the favorable relationship between the high-fidelity and low-fidelity descent characteristics. However, a straightforward implementation of a cubic polynomial response surface substitute (coded locally with assistance from the RSG [27] package) provided relatively poor models, as evidenced in Figure 4. This combination of models served a test scenario in which the descent relationship was not as favorable.



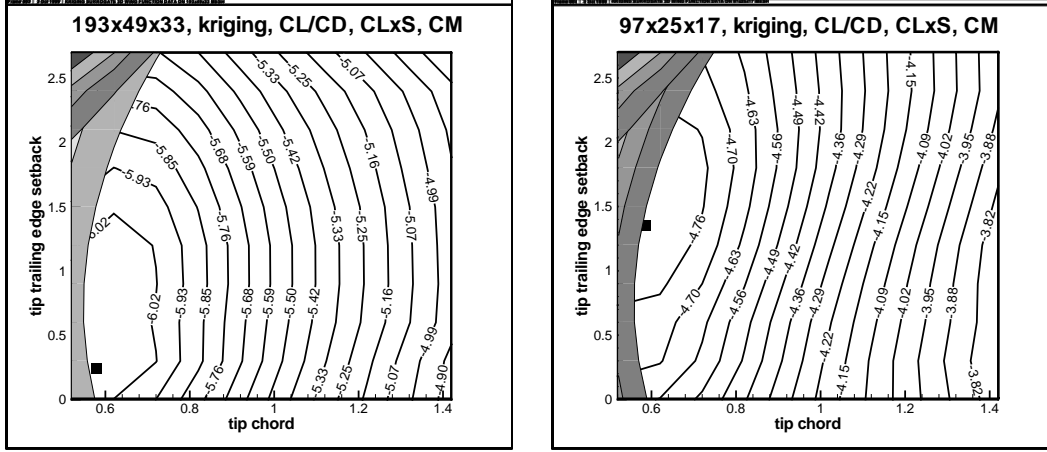


Figure 3: High-fidelity vs. low-fidelity objectives and active constraints: level sets of kriging approximation

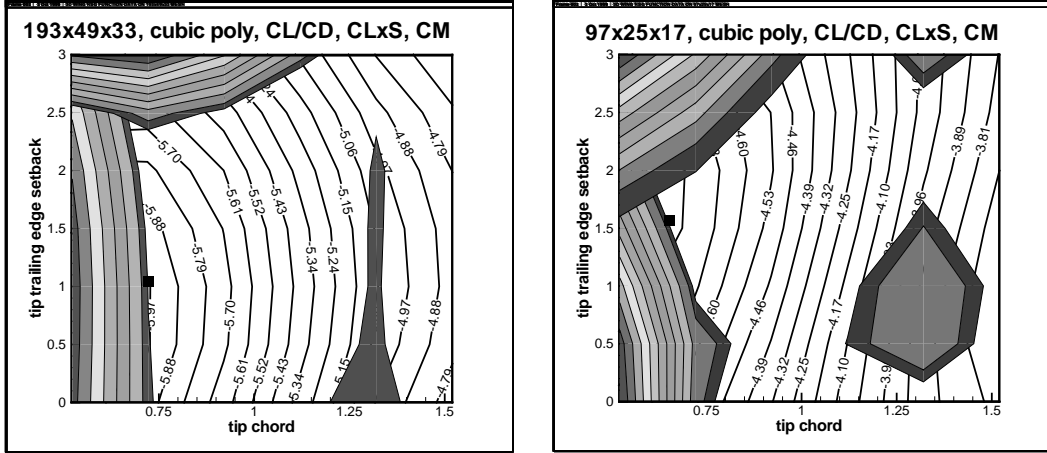


Figure 4: High-fidelity vs. low-fidelity objectives and active constraints: level sets of cubic polynomial approximation

	Full CFD analysis	Kriging	Spline	Polynomial
Augmented Lagrangian AMF		3.0 / 2.6		
SQP AMF			2.8 / 2.8	3.0 / 3.0
MAESTRO AMF	1.9 / 1.9			

Table 1: Wing optimization problem: Summary of improvement factor due to the AMF in function (first number) and sensitivity (second number) computations.

tional optimization, applied to a cubic polynomial substitute for the CFD analysis, required 31 high-fidelity functions and 31 high-fidelity sensitivities. Optimization using the SQP-based AMMO required 4 high-fidelity functions and 51 low-fidelity functions, for a total of $4 + 51/8 = 10 \frac{3}{8}$ equivalent high-fidelity functions and as many sensitivities. For a spline substitute for the CFD analysis, conventional optimization required 21 high-fidelity functions and as many sensitivities. The SQP-based AMMO required 4 high-fidelity functions, 4 high-fidelity sensitivities, 28 low-fidelity analyses,

and 28 low-fidelity sensitivities, or a total of $4 + 28/8 = 7\frac{1}{2}$ equivalent high-fidelity functions and as many sensitivities.

Table 1 summarizes the results for all three AMMO frameworks. Improvements in efficiency compared to non-AMMO versions of the same codes were consistent across the frameworks. We note that at the time of the testing, we made an effort not to “fine-tune” the AMMO codes, so as to have a fairer test.

Even though the three frameworks performed similarly, we were able to reach a number of conclusions about their relative applicability to engineering optimization problems. For instance, MAESTRO-based AMMO (as well as the underlying MAESTRO algorithm) was found to be more appropriate to bona fide MDO problems or to single-discipline optimization problems that exhibit block structure. The nature of the problems under investigation common to many single-discipline engineering optimization problems, dictated that the objective and constraint evaluations can be obtained only as a result of analysis and not be available in a modular fashion, on demand. For single-discipline problems with the objective and constraints available strictly via an analysis, SQP-based AMMO was found to be the most appropriate of the three. Surprisingly, even though the augmented Lagrangian-based AMMO was anticipated as the easiest to implement, the interplay among the algorithm’s components proved to be rather involved.

In summary, in the experiments with variable-resolution models, AMMO frameworks were found to yield consistent savings in terms of high-fidelity function and derivative evaluations compared to direct optimization without model management. We believe that greater savings can be achieved, because there is much room for improvement in the interaction among the components of the frameworks. In particular, currently the inner subproblem of minimizing the low-fidelity model is probably being solved to an unnecessarily high degree of accuracy. This can result in a degradation of overall efficiency if the low-fidelity computational expense is not negligible. Moreover, this can result in trial steps that are too long insofar as they leave the region where the trends in the corrected approximations match those of the high-fidelity responses. This can lead to candidate new designs that are not acceptable when checked using the high-fidelity calculations, and we pay the computational price for these fruitless designs. We are currently investigating this issue further.

Approximately two-fold savings were obtained for another problem, aerodynamic optimization of a 2D airfoil with variable-resolution models. Details of both demonstrations can be found in [6].

5.2 Managing variable-fidelity physics models

AMMO approaches could suffer if the lower-fidelity model does not predict the trends of the higher-fidelity model adequately. In this case, AMMO will take only a short step using the low-fidelity calculations before requiring recourse to the high-fidelity calculations, which, in effect, means optimization with high-fidelity models. Thus, in the worst case, the AMMO approach reverts to conventional optimization with the high-fidelity model.

We would anticipate seeing this worst-case scenario in variable-fidelity physics models. Although one might expect to see similar global trends, one might also expect drastically different behavior in regions where, say, in aerodynamics, viscous and shear effects become active. In such regions, the low-fidelity physics models not only fail to predict the function information accurately in a quantitative sense, but they also exhibit different trends.

To test the limitations of the AMMO approach, we have considered aerodynamic optimization of a multi-element airfoil designed to operate in transonic conditions [42]. The transonic free-stream Mach number and the multi-element nature of the airfoil makes inclusion of the viscous effects very important for obtaining physically correct results. This is confirmed in Figure 5, which depicts the Mach number in the flow for the high-fidelity and low-fidelity models for this problem, where the

boundary and shear layers are clearly visible in the viscous case. To capture the viscous effects, the

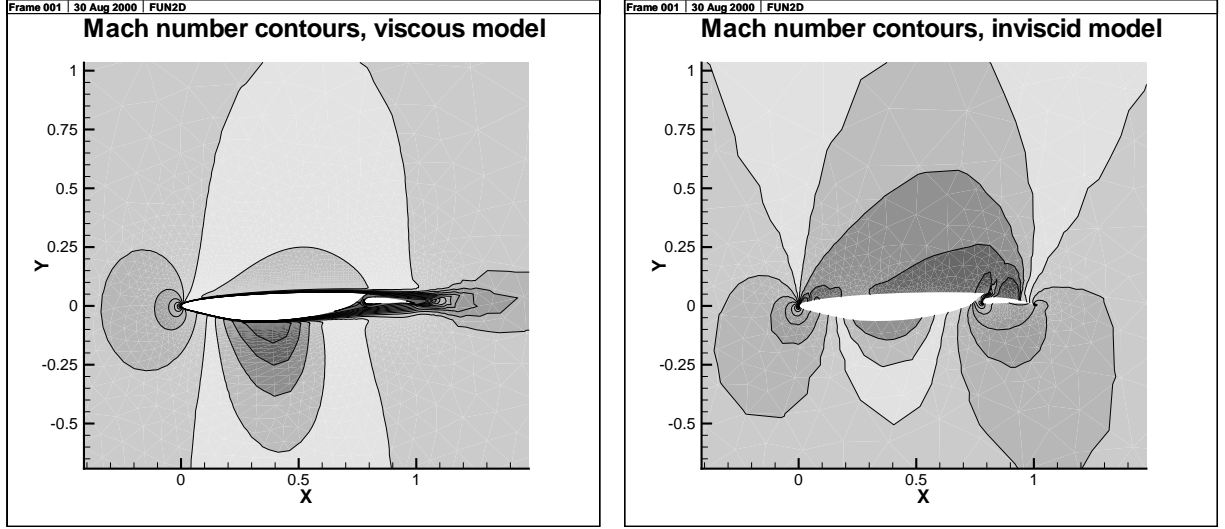


Figure 5: Mach number contours for viscous vs. inviscid model

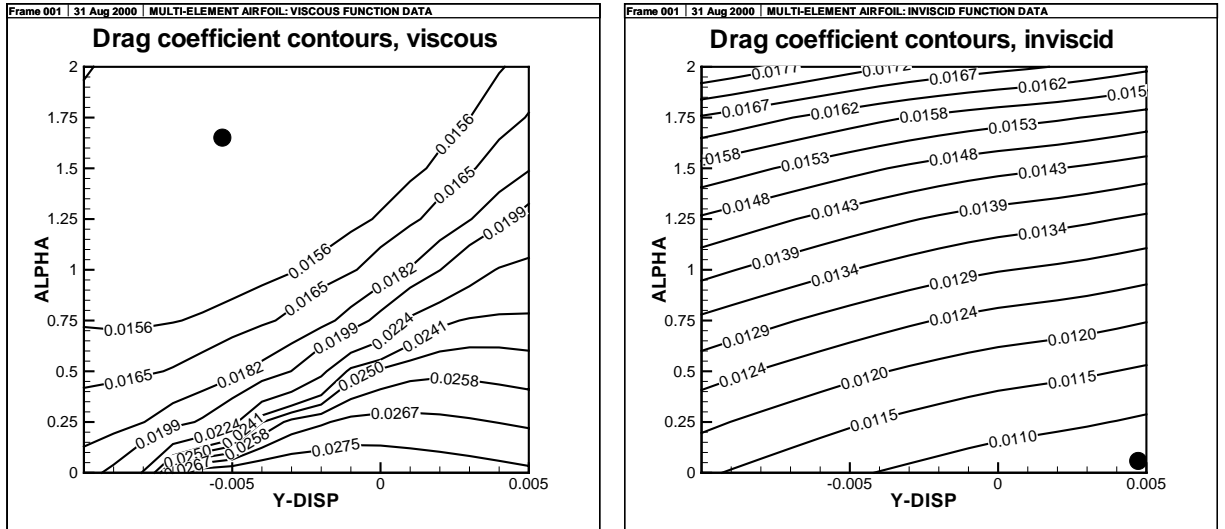


Figure 6: Drag coefficient contours of the viscous and inviscid models

governing equations of the high-fidelity model are the Reynolds-averaged Navier-Stokes (RANS) equations,

$$A \frac{\partial Q}{\partial t} + \oint_{\partial\Omega} \vec{F}_i \cdot \hat{n} \, dl - \oint_{\partial\Omega} \vec{F}_v \cdot \hat{n} \, dl = 0,$$

where \vec{F}_i and \vec{F}_v are the inviscid and viscous fluxes, respectively. Detailed description of the equation components can be found in [8,9]. The low-fidelity model is represented by the Euler equations. The analysis in both the RANS and Euler modes is provided by the flow solver FUN2D, an unstructured mesh flow solver [8]. The sensitivity derivatives are computed via a hand-coded adjoint approach [9].

The mesh for the viscous model consists of 10449 nodes and 20900 triangles. The mesh for the inviscid model comprises 1947 nodes and 3896 triangles. The free-stream Mach number is $M_\infty = 0.75$, the Reynolds number is $Re = 9 \times 10^6$, and the angle of attack is $\alpha = 1^\circ$.

Test	hi-fi eval	lo-fi eval	total t	factor
PORT with hi-fi analyses, 2 var	14/13		≈ 12 hrs	
AMMO, 2 var	3/3	19/9	≈ 2.41 hrs	≈ 5
PORT with hi-fi analyses, 84 var	19/19		≈ 35 hrs	
AMMO, 84 var	4/4	23/8	≈ 7.2 hrs	≈ 5

Table 2: AMMO performance vs PORT

Because of the importance of the viscous effects in this problem, the use of the inviscid equations for the low-fidelity model should present an important test for the present approach.

The objective of this problem is to minimize the drag coefficient by adjusting the global angle of attack and the y-displacement of the flap. The first test case is restricted to two design variables to enable visualization. The baseline case for both models was constructed at $\alpha = 1^\circ$ and zero y-displacement of the flap.

Figure 6 depicts the level sets of the drag coefficient for the viscous and inviscid models. The solution for each problem is marked with a circle. The problem manifests the most adverse situation: not only is the low-fidelity model not a good approximation of the high-fidelity model but, in fact, the descent trends in the two models are reversed. Thus the problem provides a good test of the methodology.

The computational expense necessary to calculate functions and derivatives in the viscous case is considerably greater than that for the inviscid model. We conducted our experiments on an SGITM OriginTM 2000 workstation with four MIPS RISC R10000 processors. One low-fidelity analysis took approximately 23 seconds and one low-fidelity sensitivity analysis took between 70 and 100 seconds. In contrast, one high-fidelity analysis took approximately 21 minutes and one high-fidelity sensitivity analysis took between 21 and 42 minutes to compute. The measures were taken in CPU time. Thus, the time per low-fidelity evaluation may be considered negligible compared to that required for a high-fidelity evaluation.

The problem with two variables was first solved with single-fidelity models alone using [26], in order to obtain a baseline number of function evaluations or iterations to find an optimum. The problems were then solved with AMMO. Again, for each experiment, the performance of AMMO was evaluated in terms of the absolute number of calls to the high and low-fidelity function and sensitivity calculations. Because the time for low-fidelity computations was negligible in comparison to the high-fidelity computations, we estimated the savings strictly in terms of high-fidelity evaluations.

Because the ability to handle high-dimensional problems is an important attribute of first-order AMMO frameworks and because some numerical effects could be an artifact of low dimensionality, the experiments were repeated for the problems with 84 variables. The additional variables were represented by parameters describing the shape of the airfoil. The time-per-analysis for the 84-variable case was approximately the same as for the two-variables case. However, the time-per-sensitivity increased to nearly two hours per sensitivity calculation. Table 2 summarizes the number of function (first number) and derivative (second number) computations expended in PORT and in AMMO for both problem formulations.

Given the dissimilarity between the high-fidelity and low-fidelity model, we were initially sur-

prised to find that the AMMO performed well: it consistently yielded approximately five-fold savings in terms of high-fidelity computations. Further analysis of the results revealed that the savings were not so surprising after all. For our combination of models, the β -correction worked extremely well. This is illustrated in Figure 7 for the two-variable case. The plot on the left shows the level sets of the high-fidelity model with the solution. The plot on the right depicts the level

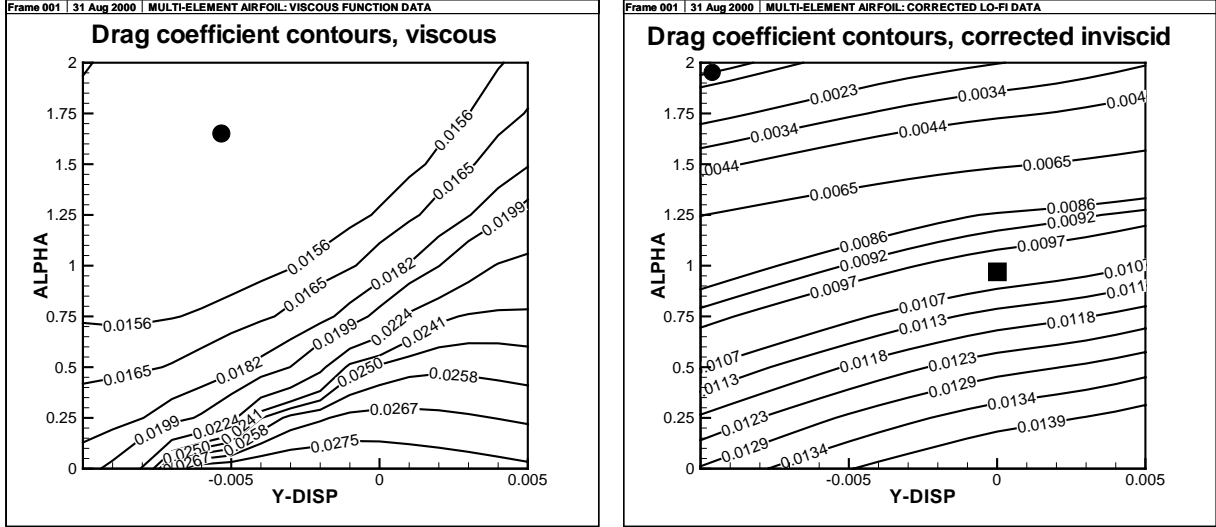


Figure 7: Drag coefficient level sets of the viscous and corrected inviscid models

sets of the low-fidelity model β -corrected at the initial point. The initial point is marked by a square. Here we have applied the correction to the entire region to visualize the effect of the correction on the low-fidelity function. The figure clearly shows that the correction, using the function and derivative information at the anchor point (at this iteration, the initial point), reversed the trend of the low-fidelity model, allowing the optimizer to find the next iterate in the left upper corner of the plot, marked by a circle. Similar analysis can be conducted for all iterations. In fact, AMMO located the solution ($\alpha = 1.6305^\circ$, flap y -displacement = -0.0048) of the high-fidelity problem already at the next iteration. The high-fidelity drag coefficient at the initial point was $C_D^{\text{initial}} = 0.0171$, the high-fidelity drag coefficient at the solution was $C_D^{\text{final}} = 0.0148$, a decrease of approximately 13.45%. The drag reduction in the 84-variable case was approximately 25%. Details of the two-variable demonstrations may be found in [7].

6 First-order vs zeroth-order model management

The first-order model management framework uses design sensitivities to good effect. However, one could alternatively consider a model management scheme based purely on response (objective and constraint) values, without recourse to sensitivities—a zeroth-order approach, as in Response Surface Methodology (see, for instance, [15,28]). We briefly compare first-order and zeroth-order model management schemes.

In classical Response Surface Methodology, one constructs approximations by building a low-order (typically, a quadratic) polynomial via regression on values of the high-fidelity responses. Modern elaborations include kriging and spline approximations constructed from high-fidelity data. Given such a model, one applies optimization to it. At some point the model is updated with new high-fidelity information, and the process repeats.

In the setting of model management, one must verify that such a zeroth-order scheme yields improvement in the true responses, just as in the first-order approach. In the first-order approach, the first-order consistency condition guarantees that if necessary, we can insure improvement using the approximation by taking very small steps. However, no such guarantee obtains in a zeroth-order model management scheme. Instead, one may have to revert to optimization on the high-fidelity model in order to make progress.

In addition, since one does not have sensitivity estimates, one cannot enforce the usual sorts of sufficient decrease criteria (e.g., the Armijo–Goldstein–Wolfe conditions in line-search methods for unconstrained minimization). The absence of information about sensitivities restricts the class of optimization algorithms one can use if one wishes to be assured of convergence of the overall process.

As noted previously, one limitation of zeroth-order approximation is the difficulty of constructing suitably useful approximations when there are more than a relatively small number of design variables (say, 20–40). In general, considerable numbers of samples of the high-fidelity response are required to adequately capture the trends in even relatively simple responses and even for a relatively small number of variables.

Moreover, there are subtleties with sampling that arise in problems of engineering interest. Imagine that the design variables are shape parameters for an airfoil, say, spline knots, as in the examples discussed previously. Our goal is to approximate (and minimize) transonic drag. Suppose we apply conventional experimental design techniques to sampling the design space in order to collect data for an approximation of the true drag. There is no guarantee that these sampled designs will correspond to physically meaningful airfoils. Instead, we will frequently compute the drag at values of the design variables that correspond to airfoils with bumps and divots and very bad aerodynamic behavior. Thus, we will build our approximation from responses at bad designs. These values will not be particularly useful in predicting the behavior of the smoother, desirable designs, because of the significant nonlinearity of the dependence of drag on airfoil shape. If we try to ameliorate this problem in sampling by confining our sampling to the span of some set of nice shapes—a reduced basis approach—we may unduly limit our design degrees of freedom.

One can use, say, variable-accuracy and variable-resolution models as predictors for promising designs in a zeroth-order approach. However, in the variable-fidelity physics example discussed in Section 5.2, the inviscid approximation has entirely the wrong trend compared to the true, viscous calculation, as Figure 6 makes clear. It cannot be used without some manner of correction as a predictor in a zeroth-order model management scheme.

7 Concluding remarks

Sensitivity information is becoming increasingly available with analysis codes. Automatic differentiation techniques make it relatively simple to generate sensitivity information. Appreciation of the utility of sensitivities in design and optimization have also led to a more widespread implementation of efficient adjoint techniques. With derivatives more readily available, and because sensitivity information is so useful in constructing approximations for the purposes of optimization, we believe the first-order framework we have described will prove very helpful.

The results obtained in our studies of AMMO and variable-fidelity models are promising. We have observed significant computational savings over conventional optimization, in terms of high-fidelity simulations. Despite the sometimes significant dissimilarity between high- and low-fidelity models, we were able to capture the descent behavior of the high-fidelity model with the assistance of the first-order correction. When the models are greatly dissimilar, first-order information appears

indispensable in obtaining reasonable descent directions. The first-order corrected approximation then leads to efficient optimization.

Given these results, we are cautiously optimistic about several much larger single-discipline test cases, multidisciplinary optimization cases, and a variety of alternative modeling options that are currently under investigation.

References

- [1] N. M. ALEXANDROV, *Multilevel and multiobjective optimization in multidisciplinary design*, in Proceedings of the Sixth AIAA/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, AIAA, 1996. AIAA Paper 96-4122.
- [2] ———, *A trust region framework for managing approximation models in engineering optimization*, in Proceedings of the Sixth AIAA/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, AIAA, 1996. AIAA-96-4102.
- [3] ———, *On managing the use of surrogates in general nonlinear optimization*, in Proceedings of the Seventh AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, AIAA, 1998. AIAA Paper 98-4798.
- [4] N. M. ALEXANDROV AND M. Y. HUSSAINI, eds., *Multidisciplinary Design Optimization: State of the Art*, Philadelphia, 1997, Society for Industrial and Applied Mathematics.
- [5] N. M. ALEXANDROV AND R. M. LEWIS, *First-order approximation/model management frameworks: analysis and performance*, December 2000. In preparation.
- [6] N. M. ALEXANDROV, R. M. LEWIS, C. R. GUMBERT, L. L. GREEN, AND P. A. NEWMAN, *Optimization with variable-fidelity models applied to wing design*. AIAA-2000-0841, January 2000. Proceedings of the 38th Aerospace Sciences Meeting and Exhibit. Accepted for publication in AIAA Journal of Aircraft. Also available as ICASE technical report 99-49.
- [7] N. M. ALEXANDROV, E. J. NIELSEN, R. M. LEWIS, AND W. K. ANDERSON, *First-order model management with variable-fidelity physics applied to multi-element airfoil optimization*. AIAA paper 2000-4886, 2000.
- [8] W. K. ANDERSON AND D. L. BONHAUS, *An implicit upwind algorithm for computing turbulent flows on unstructured grids*, Computers and Fluids, 23 (1994), pp. 1–21.
- [9] ———, *Aerodynamic design on unstructured grids for turbulent flows*, Tech. Rep. NASA TM 112867, NASA Langley Research Center, June 1997.
- [10] M. AVRIEL AND A. C. WILLIAMS, *Complementary geometric programming*, SIAM Journal on Applied Mathematics, 19 (1970), pp. 125–141.
- [11] J.-F. M. BARTHELEMY AND R. T. HAFTKA, *Approximation concepts for optimum structural design—a review*, Structural Optimization, 5 (1993), pp. 129–144.
- [12] C. BISCHOF, A. CARLE, P. KHADEMI, AND A. MAUER, *The ADIFOR 2.0 system for the automatic differentiation of Fortran 77 programs*, IEEE Computational Science & Engineering, 3 (1996), pp. 18–32.

- [13] C. BISCHOF, W. T. JONES, J. SAMAREH-ABOLHASSANI, AND A. MAUER, *Experiences with the application of the ADIC automatic differentiation tool to the CSCMDO 3-D volume grid generation code*, January 1996. AIAA Paper 96-0716.
- [14] F. G. BLOTTNER AND A. R. LOPEZ, *Determination of solution accuracy of numerical schemes as part of code and calculation verification*, Tech. Rep. SAND98-222, Sandia National Laboratories, October 1998.
- [15] G. E. P. BOX AND N. R. DRAPER, *Empirical Model-Building and Response Surfaces*, John Wiley & Sons, 1987.
- [16] V. BRAIBANT AND C. FLEURY, *An approximation-concepts approach to shape optimal design*, Computer Methods in Applied Mechanics and Engineering, 53 (1985), pp. 119–148.
- [17] S. BURGEE, A. GIUNTA, V. BALABANOV, B. GROSSMAN, W. MASON, R. NARDUCCI, R. HAFTKA, AND L. WATSON, *A coarse-grained parallel variable-complexity multidisciplinary optimization problem*, The International Journal of Supercomputing Applications and High Performance Computing, 10 (1996), pp. 269–299.
- [18] R. G. CARTER, *Numerical experience with a class of algorithms for nonlinear optimization using inexact function and gradient information*, SIAM Journal on Scientific Computing, 14 (1993), pp. 368–388.
- [19] K. J. CHANG, R. T. HAFTKA, G. L. GILES, AND P.-J. KAO, *Sensitivity-based scaling for approximating structural response*, Journal of Aircraft, 30 (1993), pp. 283–288.
- [20] A. R. CONN, N. I. M. GOULD, AND P. L. TOINT, *A globally convergent augmented Lagrangian algorithm for optimization with general constraints and simple bounds*, SIAM Journal on Numerical Analysis, 28 (1991), pp. 545–572.
- [21] P. E. GILL, G. H. GOLUB, W. MURRAY, AND M. H. WRIGHT, *User's guide for SOL/NPSOL: a Fortran package for nonlinear programming*, Department of Operations Research, Stanford University, Stanford, CA, 1983.
- [22] P. E. GILL, W. MURRAY, AND M. H. WRIGHT, *Practical Optimization*, Academic Press, London, 1981.
- [23] A. GIUNTA, *Aircraft Multidisciplinary Optimization Using Design of Experiments Theory and Response Surface Modeling Methods*, PhD thesis, Department of Aerospace and Ocean Engineering, Virginia Tech, Blacksburg, VA, 1997.
- [24] P. HAJELA, *Geometric programming strategies in large-scale structural synthesis*, AIAA Journal, 24 (1986), pp. 1173–1178.
- [25] A. JAMESON, *Essential elements of computational algorithms for aerodynamic analysis and design*, Tech. Rep. 97-68, Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, Virginia, December 1997.
- [26] L. KAUFMAN AND D. GAY, *PORT Library: Optimization and Mathematical Programming*, Bell Laboratories, May 1997.
- [27] M. KAUFMAN, *RSG: Response Surface Generation Program, Version 1.32*, Virginia Tech, January 1996. electronic version.

- [28] A. I. KHURI AND J. A. CORNELL, *Response Surfaces: Designs and Analyses*, Marcel Dekker, New York, 1987.
- [29] R. M. LEWIS, *A trust region framework for managing approximation models in engineering optimization*, in Proceedings of the Sixth AIAA/NASA/ISSMO Symposium on Multidisciplinary Analysis and Design, September 1996. AIAA Paper 96-4101.
- [30] R. M. LEWIS AND S. G. NASH, *A multigrid approach to the optimization of systems governed by differential equations*. AIAA Paper 2000-4890, 2000.
- [31] J. J. MORÉ, *Recent developments in algorithms and software for trust region methods*, in Mathematical Programming. The State of the Art, Bonn 1982, A. Bachem, M. Grötschel, and B. Korte, eds., Springer-Verlag, Berlin, 1983, pp. 258–287.
- [32] A. J. MORRIS, *Approximation and complementary geometric programming*, SIAM Journal on Applied Mathematics, 23 (1972), pp. 527–531.
- [33] S. G. NASH AND A. SOFER, *Linear and Nonlinear Programming*, McGraw–Hill, New York, 1996.
- [34] J. C. NEWMAN, III, A. C. TAYLOR, III, R. W. BARNWELL, P. A. NEWMAN, AND G. J.-W. HOU, *Overview of sensitivity analysis and shape optimization for complex aerodynamic configurations*, Journal of Aircraft, 36 (1999), pp. 87–96.
- [35] V. H. NGUYEN, J. STRODIOT, AND C. FLEURY, *A mathematical convergence analysis of the convex linearization method for engineering design optimization*, Engineering Optimization, 11 (1987), pp. 195–216.
- [36] C. RUMSEY, R. BEIDRON, AND J. THOMAS, *CFL3D: Its history and some recent applications*. NASA TM 112861, 1997.
- [37] R. E. SMITH, M. I. G. BLOOR, M. J. WILSON, AND A. T. THOMAS, *Rapid airplane parametric input design (RAPID)*, in Proceedings of the 12th AIAA Computational Fluid Dynamics Conference, San Diego, June 1985. AIAA Paper 95-1687.
- [38] J. SOBIESZCZANSKI-SOBIESKI AND R. T. HAFTKA, *Multidisciplinary aerospace design optimization: Survey of recent developments*, Structural Optimization, 14 (1997), pp. 1–23.
- [39] A. C. TAYLOR III, A. OLOSO, AND J. C. NEWMAN III, *CFL3D.ADII (version 2.0): An efficient, accurate, general-purpose code for flow shape-sensitivity analysis*. AIAA Paper 97-2204, 1997.
- [40] J. WALSH, J. TOWNSEND, A. SALAS, J. SAMAREH, V. MUKHOPADHYAY, AND BARTHELEMY, *Multidisciplinary high-fidelity analysis and optimization of aerospace vehicles, part 1: Formulation*. AIAA Paper 2000-0418, 38th Aerospace Sciences Meeting and Exhibit, Reno, NV, Jan. 10-13, 2000.
- [41] J. WALSH, R. WESTON, J. SAMAREH, B. MASON, L. GREEN, AND R. BIEDRON, *Multidisciplinary high-fidelity analysis and optimization of aerospace vehicles, part 2: Preliminary results*. AIAA Paper 2000-0419, 38th Aerospace Sciences Meeting and Exhibit, Reno, NV, Jan. 10-13, 2000.
- [42] R. T. WHITCOMB AND L. R. CLARK, *An airfoil shape for efficient flight at supercritical mach numbers*, Tech. Rep. NASA TM-X-1109, NASA Langley Research Center, 1965.